

АВТОМАТИЧЕСКИЙ ПОДБОР ГИПЕРПАРАМЕТРОВ НЕЙРОННОЙ СЕТИ МЕТОДОМ КРИГИНГА

Тя-Шен-Тин Е. Н.^{1*}, Разумов А. А.^{1,2}, Ушенин К. С.^{1,2}

¹)Уральский федеральный университет имени первого Президента России Б. Н. Ельцина, г. Екатеринбург, РФ

²)Институт иммунологии и физиологии УрО РАН, г. Екатеринбург, Россия

*E-mail: tyashentin@yandex.ru

FITTING OF DEEP NEURAL NETWORK HYPERPARAMETERS BY KRIGING

Tya-Shen-Tin Y.N.^{1*}, Razumov A.A.^{1,2}, Ushenin K.S.^{1,2}

¹)Ural Federal University, Yekaterinburg, Russia

²)Institute of Immunology and Physiology of the Russian Academy of Sciences, Yekaterinburg, Russia

A choice of optimal hyperparameters requires during a deep neural network training for reaching of the loss function minimum value on a test or validation dataset. However, it is possible to automate this process by algorithms minimizing the value of a loss function. In this paper we perform searching of hyperparameters for deep neural networks using kriging, that is a regression based on Gaussian processes.

При обучении глубоких нейронных сетей необходимо подобрать оптимальные значения гиперпараметров. Как правило, данная задача решается экспертом. В данной работе процесс поиска оптимальных параметров автоматизирован при помощи алгоритма минимизации функции - кригинга (регрессии на основе гауссовских процессов) [1].

Метод кригинга осуществляет приближение минимизируемой функции и обеспечивает нахождение приближенного глобального минимума за меньшее число вызовов функции, в сравнении с альтернативными классами методов глобальной оптимизации с ограничениями.

Оптимизация таких гиперпараметров как количество нейронов и размер мини-выборки выполнялось для нейронных сетей типа “многослойный персептрон” и “сверточная нейронная сеть”, решающих задачу распознавания рукописных цифр из базы данных MNIST[2]. Аппроксимированное минимальное значение функции потерь было найдено за 20 вызовов функции.

1. Forrester A., Sobester A., Keane A. Engineering design via surrogate modelling: a practical guide. – John Wiley & Sons, 2008.
2. LeCun Y. The MNIST database of handwritten digits / <http://yann.lecun.com/exdb/mnist/> – 1998.